Mediating and Analyzing Social Data

Ying Ding[1], Ioan Toma[2], Sin-Jae Kang[3], Zhixiong Zhang[4], Michael Fried[2]

[1] Indiana University, 1320 E 10th, Bloomington, IN 47405 USA
dingying@indiana.edu
[2]University of Innsbruck, Innsbruck, Austria
{ioan.toma; michael.fried}@sti2.at
[3]Daegu University, Daegu, South Korea
sjkang @daegu.ac.kr
[4]Library of Chinese Academy of Sciences, Beijing, China
zhangzx@mail.las.ac.cn

Abstract. Web 2.0 is turning current Web into social platform for knowing people and sharing information. The Web is strongly socially linked than ever. This paper takes major social tagging systems as examples, namely delicious, flickr and youtube, to analyze the social phenomena in the Social Web in order to identify the way of mediating and linking social data. A simple Upper Tag Ontology (UTO) is proposed to integrate different social tagging data and mediate and link with other related social metadata.

Keywords: Social Tagging, data mediation, Social Web, ontology

1 Introduction

Web 2.0 is turning current Web into social platform for knowing people and sharing information. The Web is strongly socially linked than ever. The term "Social Web" was introduced in 1998 by Peter Hoschka [1] who tried to stress the social medium function of the Web. From Wikipedia, the Social Web is defined as an open global distributed data sharing network which links people, organizations and concepts. Current Web 2.0 is the main stream of the Social Web which provides platform and technologies (such as wiki, blog, tag, RSS feed, etc.) for online collaboration and communication.

The online publishing in Web 2.0 made everything so easy that anyone who can write or type can publish their data to the Web. This revolution significantly stimulates the amount of normal users to get involved to the Web communication; those of them are just teenagers or old people. One of the new ways of adding data to the current Web is tagging which reflects community effort on organizing and sharing information. Tagging is a kind of adding keywords through typed hyperlinks. Now the web is changing from hyperlinked documents to typed hyperlinked data web.

As from current Web 2.0, we already evident human-created metadata (such as tags) which are growing daily on the Web. This trend will further lead to more similar metadata as well as metadata generated from Semantic Web community which is

ontologically explicitly defined, for example, FOAF (metadata for friends), SKOS (metadata for taxonomies), DOAP (metadata for project), RSS (metadata for news), SIOC (metadata for social networks), Dublin Core (metadata for documents), GEO (metadata for geographic coordinates), GeneOnt (metadata for human genes), microformat (metadata for Social Web) and so on.

Furthermore, machine can also start to contribute data to the Web as machine can generate data automatically based on pre-defined ontologies. Those metadata and data are not isolated but interlinked. Based on four principles of linking open data proposed by Tim Berners-Lee, more and more linked semantic data are available (see Link Open Data initiative[1]). Those kind of linking is mainly through owl:sameAs or fofa:knows to link different concepts or instances. We call those links semantic links. These powerful semantic links will weave the current Web to its future. The future Web is the Web of semantically linked semantic data.

This paper takes major social tagging systems as examples, namely delicious, flickr and youtube, to analyze the social phenomena in the Social Web in order to identify the way of mediating and linking social data. The main contributions of our work include:

- Modeling social tagging data according to proposed Upper Tag Ontology (UTO).
- Linking UTO with other related social metadata (such as FOAF, DC, SIOC, SKOS, etc.)
- Crawling tag data from major social tagging systems and integrating them according to UTO.
- Clustering crawled tagging data.

According to above, this paper is organized as follows. Section 2 gives the detailed description of how to model social tagging data, how to link them with related social semantics, how to crawl social tagging data and how to analyze tagging data via clustering. Section 3 discusses the related work. Section 4 concludes the paper and presents some future work.

2 Social Tagging

Tag is a keyword used to categorize online objects. The goal of tagging is to make a body of information increasingly easier to search, discover, share and navigate over time. Social tagging is not simply just tagging, tags are social metadata generated from collective intelligence. The consensus of tags forms social semantics which are called folksonomies. It is bottom-up approach and reflects collective agreement. It speaks the same language as the users and makes the things easier to find.

[1] http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData

2.1 Modelling social tagging data

We can tag bookmarks (del.icio.us), photos (flickr), videos (YouTube), books (LibraryThing), Music (Last.fm), citations (CiteULike), blogs (Technorati), etc. Tag is nothing special than a typed hyperlink. We can use "rel" attribute to create typed hyperlink. There are many social networks providing tagging services, here we take three major social tagging systems, namely delicious, flickr, and youtube, to analyze their social tagging behavior. Based on this analysis, we propose Upper Tag Ontology (UTO) which is originated from Tag Ontology proposed by Tom Gruber [2]. In his tag ontology, he proposed five key concepts which are object, tag, tagger, source and vote. Here in UTO, we add another three concepts: comment, date and tagging. Because most of the social networks contain information about comments for the tags or objects, these provide extra information for us to better understand the meaning of the tags or objects. Date is another important concept for us as it depicts the evolution of the tags and tagging behavior. It can also help us to unveil the hidden social changes inside a social network. The tagging concept plays a role to interlink all these core concepts together. Itself does not have real meaning. Furthermore, we add has_relatedTag relationship to tag concept itself. More details about modeling social tagging data were discussed in [3].

Let O be UTO ontology,

$$\mathrm{O} = (C, \Re) \tag{1}$$

Where $C = \{c_i, i \in N\}$ is a finite set of concepts

$\Re = \{(c_i, c_k), i, k \in N\}$ is a finite set of relations established among concepts in C.

In UTO,

$$C = \begin{Bmatrix} Tag, Tagging, Object, Tagger, Source, \\ Date, Comment, Vote \end{Bmatrix},$$

$$\Re = \begin{Bmatrix} has\_relatedTag, has\_tag, has\_object, \\ has\_source, has\_date, has\_creator, \\ has\_comment, has\_vote \end{Bmatrix}$$

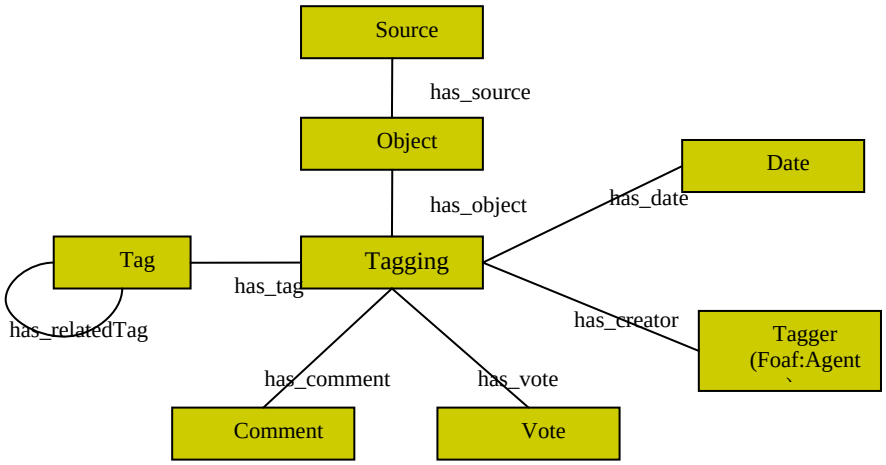

Fig. 1. Upper Tag Ontology (UTO)

Figure 1 presents the concepts and relations of UTO. As we see, UTO is a very small and simple ontology with 8 concepts and 8 relationships (see Table 1 and Table 2). The tagging concept acts as a virtual connection among different concepts in UTO. It does not have real meaning rather than the function of linking some core concepts. For instance, it is hard to tell whether the date is for tag or the tagging behavior, or comment can be viewed as being added to tag or to object directly. So most of the relations in UTO are defined as transitive so that comment can be connected to object via tagging or to tag via tagging.

Table 1. Concepts in UTO

Concept	Synonyms	Description	Value Type	Instance
Tagging		Tagging is the concept which is created to link other concepts. It, itself, does not have any real meaning.	string	e.g., tagging
Tag	keyword	Tag is the keyword which users add to object	string	e.g., design, web2.0, instructional_design , tutorials
Tagger	user	Tagger is the user who tags object	string	e.g., sborrelli
Object	Online object	Object is the thing which tagger is tagging. It can be bookmarks (URLs), photos, videos, musics, books, slides, etc.	string	e.g., www.commoncraft. com/show
Source	Social network	Source is the place where the object is hosted. It can be del.icio.us, flickr, youtube, etc.	string	e.g., del.icio.us

Comment	note	Comment is what the tagger adds to the object or tag during the tagging.	string	e.g.,The CommonCraft Show 1 Common Craft – Social Design for the Web.
Date	time	Date is the time stamp of the tagging behavior. Format is "Mmm JJ".	date	e.g., Jun 07
Vote	favorite	Tagging can be viewed as voting. Vote can be the number of different taggers tagging this bookmark (del.icio.us), a photo been favored (flickr), or a video been voted (youtube)	integer	e.g., 103 (there are 103 taggers tagged this bookmark)

According to formula (1), when $r \in \Re$, $i \in \mathrm{I}$ (I is the instances of ontology O), $h, j, k \in N$

$r^{'}$ is the inverse relation of r, when $i_j, i_k \in \mathrm{I}$, then $r(i_j) = i_k \Rightarrow r^{'}(i_k) = i_j$

r is transitive, when $i_h, i_j, i_k \in \mathrm{I}$, then $r(i_h) = i_j, r(i_j) = i_k \Rightarrow r(i_h) = i_k$

r is symmetric, when $i_j, i_k \in \mathrm{I}$, then $r(i_j) = i_k \Leftrightarrow r(i_k) = i_j$

Table 2. Relations in UTO

Relation	Domain	Range	Cardinality	OWL Type	Math properties	Inverse relation
has_tag	Tagging	Tag	N	Object Property	Transitive	is_tag_of
has_ relatedTag	Tag	Tag	N	Object Property	Transitive, Symmetric	--
has_creator	Tagging	Tagger	1	Object Property	--	is_creator_of
has_object	Tagging	Object	1	Object Property	--	is_object_of
has_date	Tagging	Date	1	Object Property	--	--
has_source	Object	Source	N	Object Property	--	is_source_of
has_ comment	Tagging	Comment	N	Object Property	--	is_ comment_of
has_vote	Tagging	Vote	N	Object Property	--	is_vote_of

UTO is different comparing to folksonomy which focuses on the meaning of tags. With the basic ontology design idea of “making it easy and simple to use”, UTO is designed to capture the structure of the social tagging behavior rather than the topic or meaning of the tags. It aims to model the structure of the tagging data in order to integrate different tagging data and link them with existing social metadata.

2.2 Linking social data

As mentioned previously, data should be interlinked. Link is changing from normal hyperlink in Web 1.0, to typed hyperlink in Web 2.0, till semantic link in web 3.0. First of all, we try to link documents, therefore we have linked online documents as Web 1.0. Then, we are adding more metadata to those documents and turning unstructured information into structured information. Later on, we should semantically link those structured information so as to form so called Web 3.0 or Semantic Web. Social tagging plays an important role here by not only structuring information but also linking structured data.

Table 3 shows the alignment between UTO and other social metadata, such as FOAF, DC, SIOC and SKOS. Here we try to make the alignment as simple as possible because the complicated alignment may generate problems or double the complicity of application. So here we focus mainly on class mapping with the consideration of equal and sub-class mapping. For instance, “Tagger” concept equals to foaf:Person, sioc:User, dc:Contributor and dc:Creator; it is the subclass of foaf:Agent, foaf:Group, foaf:Organization and sioc:Usergroup. “Tag” concept equals to skos:Concept; it is subclass of dc:Subject and skos:Subject. “Object” concept is superclass of foaf:Document, foaf:Image, sioc:Post, sioc:Item, dc:Text and dc:Image. “has_relatedTag” relationship is the super-property of skos:narrower, skos:broader and skos:related.

Aligning UTO with other existing social semantics enables easy data integration, mash-ups different semantics and interlinks structured data. Based on these integrated data, we can perform tag search across multiple sites, applications, sources, hosts and mine relations (associations) cross different platforms and applications. For instance, we can do the following queries: finding friends of Stefan who tagged “spicy-Chinese-food” by aligning FOAF with UTO; finding different blogs, wikis, or discussion groups which Stefan or his friends join and discuss the topic on “spicy-Chinese-food” by aligning FOAF, SOIC with UTO, etc. Associations among tag, tagger and objects can be mined as well. For instance, we can mine the social network relations of taggers through foaf:knows by aligning FOAF with UTO; we can mine the relation or association of tags through skos:broader, skos:narrower or skos:related; we can use co-occurrence technologies to mine the association among tags, taggers and objects, etc.

Table 3. Different ontology alignment with UTO

UTO	FOAF	SIOC	DC	SKOS
Tagging	--	--	--	--
Tag	--	--	$\subseteq$ Subject	= Concept $\subseteq$ Subject
Tagger	= Person $\subseteq$ Agent $\subseteq$ Group $\subseteq$ Organization	= User $\subseteq$ Usergroup	= Contributor = Creator	--
Object	$\supseteq$ Document $\supseteq$ Image	$\supseteq$ Post $\supseteq$ Item	$\supseteq$ Text $\supseteq$ Image	--
Source	--	$\subseteq$ Community	= Source	--
Comment	--	--	$\subseteq$ Description	--
Date	--	--	= Date	--
Vote	--	--	--	
has_relatedTag	--	--	--	$\supseteq$ narrower $\supseteq$ broader $\supseteq$ related

Notes: according to formula (1), $c_i, c_j \in C$, $c_i \subseteq c_j \Leftrightarrow c_i$ is the sub-class of c_j, while, $c_i \supseteq c_j \Leftrightarrow c_i$ is the super-class of c_j, while $c_i = c_j \Leftrightarrow c_i$ equals to c_j. The same is valid for relationship.

2.3 Crawling social tagging data

Social Tagging crawler (in short ST crawler) is a developed multi-crawler designed for crawling major social tagging systems including del.icio.us, flickr and youtube [4]. This crawler is based on the "Smart and Simple Webcrawler"[2] and UTO. Figure 2 shows the detailed class diagrams of the crawler.

[2] https://crawler.dev.java.net/

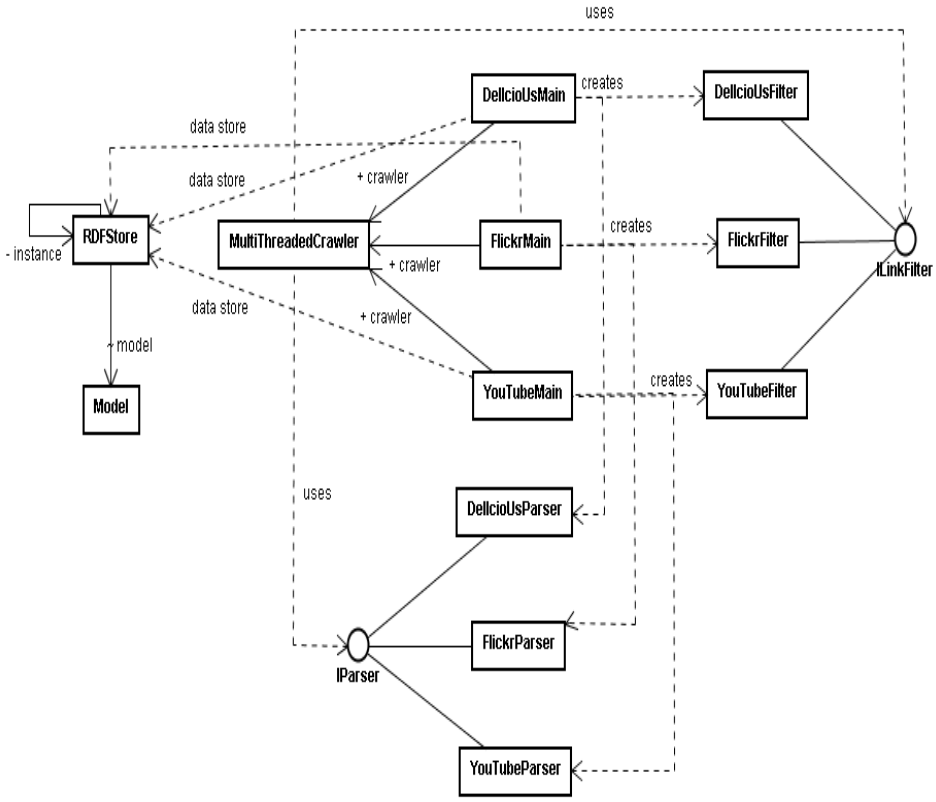

Fig. 2. Class diagram overview of the ST crawler

The ST crawler is written in Java with Eclipse IDE 3.2 on Windows XP and Ubuntu 6.04. Data has been cleaned up using linux batch commands. ST crawler can start from one or a list of links. There are two crawling models:

- Max Iterations: Crawling a web site through a limited number of links. It needs a small memory footprint and CPU usage.
- Max Depth: A simple graph model parser without recording incoming and outcoming links. It uses filter to limit the links to be crawled.

Finally, ST crawler has crawled social tagging data from delicious, flickr and youtube and modelled them according to UTO. These data are represented in RDF triples and stored in Jena. In the summer of 2007, we use ST crawler to crawl tagging data from these three websites. After one-week crawling, the crawled output contains several RDF files with a complete file size of 2.10GB. In detail:

- 16 del.icio.us data files at a size of 1.64GB
- 3 flickr data files at a size of 233MB
- 3 youtube data files at a size of 234MB

2.4 Clustering social tagging data

Based on above crawled data, we took the 1.64GB tagging data crawled from delicious as one sample to analyze social feature of its community. The crawled tagging data from delicious contains 462,733 taggers, 404,388 tags and 483,564 bookmarks. All these tag data are represented in RDF and stored in Jena. We took the tag data as they are and did perform data cleaning (for instance, stemming and checking with WordNet). By querying these data, we got the top 20 highly ranked tags and top 20 highly ranked bookmarks during that time (see Table 4).

Table 4. Top 20 highly ranked tags and bookmarks in del.icio.us

Rank	Tag	Tag Frequency	Bookmark	Bookmark Frequency
1	blog	141,871	en.wikipedia.org	26,745
2	system	120,673	www.youtube.com	14,990
3	design	109,249	community.livejournal.com	6,594
4	software	87,719	www.google.com	6,376
5	programming	83,665	www.w3.org	6,193
6	tool	83,461	news.bbc.co.uk	5,718
7	reference	74,602	www.flickr.com	5,645
8	web	70,538	java.sun.com	5,538
9	video	65,226	www.nytimes.com	5,222
10	music	61,246	www.microsoft.com	5,219
11	art	57,970	lifehacker.com	5,207
12	linux	47,965	www-128.ibm.com	4,569
13	tutorial	41,844	www.codeproject.com	4,429
14	java	40,780	www.wired.com	4,269
15	news	40,652	video.gooogle.com	4,261
16	game	39,391	www.techcrunch.com	3,818
17	free	39,006	www.bbc.co.uk	3,318
18	development	37,914	www.readwriteweb.com	3159
19	business	35,272	blogs.msdn.com	3,121
20	internet	34,580	msdn2.microsoft.com	2,950

It seems that blog topic dominates del.icio.us. Most of taggers are IT guru as system, design, software, programming, tool are ranked very high. Web and Internet are evergreen topics among the community. People like to share music, video, news, game which are popular topics in social web. People like things for “free” (as free is ranked as 17th). Highly ranked bookmarks include major social networks (youtube, livejournal, wikipedia, flickr), major news (BBC, New York Times), major computer giants (Microsoft, Google, IBM, Sun) which show the social impact of these websites.

We conduct clustering analysis based on the same data set by using X-Means algorithm. X-Means is an unsupervised clustering algorithm which one can set minimum and maximum number of clusters while training [5]. Table 5 presents some interesting clusters from our analysis.

Table 5. Tag clusters in del.icio.us

Cluster	Tags
1	ajax, c, code, development, html, java, library, net, python, rails, rudy
2	dictionary, English, language, literature, writing
3	comic, entertainment, film, forum, japan, Japanese, movie, radio, streaming, television, tv
4	calculator, conversion, convert, converter, currency, euro, exchange
5	account, bank, banking, bill, consumer, credit, deal, doctor, financial, healthcare, insurance, loan, medical, medicare, medicine, savings
6	air, apartment, building, cleaning, do, fire, guide, house, housing, move, rental, safety, studio
7	Black, blue, brown, fairy, flower, gratis, leather, line, neo, pink, red, skull, stripes, style, Sweden, Swedish, vintage, white, yellow
8	culture, history, philosophy, politics, religion
9	astronomy, earth, geography, german, map, nasa, space, world
10	font, illustration, inspiration, portfolio, typography

Cluster 1 contains 11 tags and is about programming languages. Cluster 2 has 5 tags with the topics around natural language and dictionary. Cluster 3 has 11 tags and is talking about entertainment, movie, video and radio. Cluster 4 contains 7 tags on currency conversion. Cluster 5 contains 16 tags on banking and insurance. Cluster 6 contains 13 tags on housing. Cluster 7 contains 19 tags on color. Cluster 8 on culture, Cluster 9 on geography and Cluster 10 on portfolio. Although we cannot rank clusters, comparing with Table 4 top 20 highly ranked tags, we can find out that programming languages and entertainment (video, film, movie, news and radio) are both reflected in Table 4 and Table 5. Furthermore, we can draw some interesting conclusions from Table 4 and Table 5:

- Taggers like to use adjectives (such as color) as tags to categorize their bookmarks.
- When tagging bookmarks related to currency conversion, housing and banking, taggers tend to use quite similar tags (see Cluster 4, Cluster 5 and Cluster 6)
- Two major topics in delicious are programming and entertainment. This also means that the main user groups in delicious contain users who are interested in programming and users who are interested in entertainment.

3 Related works

In 2005, Tom Gruber proposed the idea of using ontology to model tagging data. His idea has been further formalized and published in 2007 [2]. His tag ontology contains tagging (object, tag, tagger, source, + or -). He introduced vote to tag ontology and uses it for collaborative filtering. UTO contains more concepts and relations comparing to his tag ontology, such as date, source, comment, etc. Furthermore, UTO also focuses on integration with other existing social metadata in order to achieve data

integration. UTO is based on Gruber's idea and goes a bit further on ontology alignment and data integration.

SCOT[3] (Social Semantic Cloud of Tags) Ontology semantically represents the structure and semantics of a collection of tags and to represent social networks among users based on the tags. The core concepts of SCOT include Tagcloud and Tag. SCOT uses URI mechanism as unique tag namespace to link tag and resource. SCOT ontology is based on and linked to SIOC, FOAF and SKOS. It uses SIOC concepts to describe site information and relationships among site-resources. It uses FOAF concepts to represent a human or machine agent. It uses SKOS to characterize the relations between tags. While UTO does not care much of tagcloud and it is defined in such a way which can be further aligned with many other social metadata, such as DC, microformat, etc.

Holygoat Tag Ontology[4] models the relationship between an agent, an arbitrary resource and one or more tags. Taggers are linked to foaf:agents. Taggings reify the n-ary relationship between tagger, tag, resource and data. This ontology also links itself to RSS and dc, such as rss:item, rss:category, rss:pubDate, rss:link and dc:subject by using rdfs:subClassOf or rdfs:subPropertyOf. Based on these, they can perform some simple subsumption inference. This approach goes a bit deep to semantic web by utilizing ontology reasoning and inference. UTO aims to keep things simple and easy to use therefore ontology reasoning and inference is not considered at this stage.

MOAT Ontology[5] is a lightweight ontology to represent how different meanings can be related to a tag. It focuses on providing unique identifier to tag which associated semantic meaning to the tag. It is based on Holygoat Tag Ontology to define tag object. MOAT assumes that there exists a unique relationship between a tag and a label that a tag can have a unique MOAT identifier in the semantic web. UTO cares more about the structure of the tagging behavior rather than the meaning of the tags. But provide unique identifier to tag is always a helpful and important issue to social tagging and furthermore to web in general.

4 Conclusion and future work

The current Web has experienced tremendous changes to connect information, knowledge, people and intelligence. There are a couple of existing efforts trying to bring the Web to its next generation. The Semantic Web is one of the efforts embedded significantly in academic artificial intelligence area. It has the long-term vision to make the Web as the global brain of human and machine by representing data in machine understandable way and automating the mediation of data and services. Meanwhile, Web 2.0 represented Social Web has successfully motivated users to share information and collaborate each other directly via the Web [6].

Web 2.0 is not completely different from the Semantic Web [7]. As Sir Tim Berners-Lee mentioned "the Semantic Web is an extension of the current Web in

[3] http://scot-project.org/

[4] http://www.holygoat.co.uk/projects/tags/

[5] http://moat-project.org/ontology

which information is given well-defined meaning, better enabling computers and people to work in cooperation[6]". Web 2.0 not only extends the communication dimensions (publishing, commenting and arguing) but also tries to add extra contextual information (we can call it "social metadata") to the current Web data in a social and informal way (e.g. tagging, bookmarking and annotating). The power of the Semantic Web lies in the potential for interoperability through some well-defined metadata in machine understandable way and logic reasoning support [8]. Module and layer design principle in the Semantic Web (e.g. ontologies, languages and services) paves the way for reuse and intelligent search with more granularity and relevance [9]. Web 2.0 provides scalable community-powered information sharing platform, while the Semantic Web adds valuable machine understandable metadata to enable efficient and automatic way of heterogeneous information sharing and cross-portal communication and collaboration [10].

This paper takes social tagging systems as examples and aims to identify some pragmatic ways of utilizing Semantic Web and Social Web phenomena to structure unstructured information. A simple Upper Tag Ontology (UTO) is proposed to integrate social tagging data from different social networks and mediate with other related social metadata so that data are interlinked. Furthermore, the broader way of data mediation (mediate different ontological concepts or relationships) can be established based on community driven methods with the consideration of instances and contextual information. It has the following important features:

- *Community driven mediation based on collective intelligence:* Ontology mediation is one of the hardest problems in the Semantic Web which is mainly achieved formally and manually. These kinds of approaches can be hardly adopted by the Web due to the scalability issue. Social Web changes the current Web into a community platform where ordinary users participate daily for communication and collaboration. This social synergy can be used for data mediation as mediation itself is a kind of activity supporting communication and collaboration. Community driven mediation based on social collective intelligence can be an appropriate approach for data mediation. Furthermore social web services can provide further support for browsing and querying mediated data.
- *Instance-based metadata mediation:* There are already some existing researches on instance-based metadata mediation from the Semantic Web and database area. But they are more focusing on the formal transformation problem between schema and instances. Ideas on how to advance the data mining techniques to mediate metadata based on instances and contextual information around the data and metadata can be further explored. Especially, due to the Social Web effect, social involvement of the users should be significantly considered during the process and should be integrated into the approach.
- *Efficient mashing-up of Social Web services and metadata semantics:* In its current state, the Web is often described as being in the Lego phase, with all of its different parts capable of connecting to one another. Properly mashing-

[6] http://www.w3.org/2001/sw/EO/points

up social services can assist the mediation process and further enable the browsing and querying of the mediated data.

Social aspect of the Web indeed influences fundamentally the usage and sharing of the web information. The Web relies on people serving useful content, linking them and providing trust and feedback. The massive participation of the web users has significantly increased the heterogeneity of the Web. On the other hand, it has created the additional way for data integration, namely integration by collective intelligence. By tagging and sharing data, intuitively they also enrich the contextual information of the concepts and relations. Here we take social tagging systems as examples to identify some pragmatic ways of utilizing Semantic Web and Social Web phenomena to realize data mediation and integration. A simple Upper Tag Ontology (UTO) is proposed to integrate different social tagging data and mediate with other related social metadata. In the future, we would like to put some efforts to mine some associations among these tagging data in order to portray tagging behavior in current social networks. We can also build up recommender systems based on these associations. Furthermore, some efficient statistical methods can be identified to extract mediation rules based on instances and contextual information.

References

1. Hoschka, P.: CSCW research at GMD-FIT: From basic groupware to the Social Web. ACM SIGGROUP Bulletin, 19(2), 5-9 (1998).
2. Gruber, T.: Ontology of Folksonomy: A Mash-up of Apples and Oranges. International Journal on Semantic Web & Information Systems, 3(2) (2007). Available at: http://tomgruber.org/writing/ontology-of-folksonomy.htm
3. Ding, Y., Toma, I., Kang, S., Fried, M. and Yan Z.: Data Mediation and Interoperation in Social Web: Modeling, Crawling and Integrating Social Tagging Data. WWW2008 Workshop on Social Web Search and Mining (SWSM2008), April 21-25, Beijing, China (2008).
4. Fried, M.: Social Tagging Wrapper. Bachelor Thesis, Institute of Computer Sciences, University of Innsbruck, Austria (2007).
5. Pelleg, D., Moore, A. W.: X-means: Extending K-means with Efficient Estimation of the Number of Clusters. In: Seventeenth International Conference on Machine Learning, 727-734 (2000).
6. Hinchcliffe, D.: The State of Web 2.0. Web Services Journal (2006). Available: http://web2.wsj2.com/the_state_of_web_20.htm
7. Berners-Lee, T., Hendler, J. and Lassila, O.: The Semantic Web. Scientific American, 284(5):34–43 (2001).
8. Antoniou, G. and van Harmelen, F.: A Semantic Web Primer. MIT Press (ISBN 0-262-01210-3) (2004).
9. Gomez-Perez, A., Fernandez-Lopez, M. and Corcho, O.: Ontological Engineering: Advanced Information and Knowledge Processing. Springer. (2003).
10. Mika, P.: Social Networks and the Semantic Web. Springer (ISBN: 978-0-387-71000-6) (2007).